SPECIFICATION

SPEECH INTERACTIVE INTERFACE UNIT

BACKGROUND OF THE INVENTION

5

1. Field of the Invention:

The invention relates to an interactive speech interface unit for operating applications using an interactive speech.

2. Description of the Related Art:

10   A speech interface unit for operating applications by speech has been recently devised. Fig. 22 shows an example of the operation of an application using a conventional interactive speech. Although an actual input by a user and a response by a system are performed by a sound or speech, a statement comprised of the mixture of kanji and kana is used

15   hereafter for the convenience of an explanation.

In the interactive speech, flow of dialog need be controlled so as to perform the dialog smoothly. The flow of dialog means a manner of response by a system to the input by a user wherein when a system performs appropriate responses, an efficient interactive speech function can

20 be realized.

An interactive sequence means data which the system holds therein for the purpose of controlling the flow of dialog. The interactive sequence is a network describing statuses of the system in the dialog, an event (results of analysis of a user utterance, results of data, statuses of various flags), an

25   action for operating something (feedback from the system to a user, application operations, setting of various flags) and a next transitive status.

Fig. 23 is an example of a conventional interactive sequence (an interactive sequence as disclosed in a second reference, described later). Characters encircled by squares represent statuses. When an event occurs

30   at every interactive status, an action to be executed by a system relative to the event occurred and a status to be transitive after execution of the action are described. The interactive sequence is first started from an initial status and ended when it transits to an end status. Further, with the interactive sequence, in some status, that status is once stored and another

35   interactive sequence is executed, then an operation can be restarted from the stored status of a source interactive sequence upon termination of the

interactive sequence like a subroutine call in a program language.

In this case, an execution extending from a starting status to an ending status of the interactive sequence corresponding to the subroutine call becomes one action of the source interactive sequence.

5 The method of controlling flow of dialog is changed depending on an application operated by interactive speech and a field which the application handles. There are following references relating to an invention for facilitating easiness of conversion (hereinafter referred to as field conversion property) when the application or the field is changed.

10 First Reference: Japanese Patent Laid-Open Publication No. 8-77274

Second Reference: Japanese Patent Laid-Open Publication No. 11-149297

In the first reference, a module referred to as "interactive sequence switching part" selects a corresponding one interactive sequence pattern from interactive sequence patterns stored in an interactive sequence storage

15 part in response to kinds of service which is selected by a user. A field conversion property is enhanced by replacing an interactive sequence pattern to be stored in the interactive sequence storage part.

In the second reference, an interactive sequence is divided into two layers of interactive sequences wherein an upper layer is for a general part

20 and a lower layer is for a field dependent part, and wherein the lower layer interactive sequence is subjected to a subroutine call from the upper layer interactive sequence. When the lower layer interactive sequence is replaced by another interactive sequence, the field conversion property is enhanced.

25 However, in the technique as disclosed in the first reference, interactive sequences are replaceable as a whole at every field, application so that the efficiency of the preparation of the interactive sequence is not achieved. As a result, it has been necessary to develop interactive sequences at every corresponding applications and fields.

30 Further, although the field conversion property is enhanced by re-preparing only the lower layer interactive sequences in the invention as disclosed in the second reference, there is a possibility that the modification of the lower layer affects the upper layer, and hence the field conversion property is not always sufficient.

35 Still further, in either reference, there is no means for a user to customize a flow of dialog.

2

SUMMARY OF THE INVENTION

It is an object of the invention to provide an interactive speech interface unit which is high in field conversion property and is easily customized by a user.

To achieve the above object, the interactive speech interface unit of the invention comprises speech recognition means for recognizing input speech of user utterance and converting the recognized input speech into a character string, input statement analysis means for analyzing the character string and converting the analyzed character string into semantic representation, interactive control means for controlling flow of an interactive status and accessing an application, output statement generation means for generating an intermediate language to be outputted to the user, speech generation means for converting the intermediate language into speech and outputting the speech, and application interface means for accessing the application using the semantic representation outputted from the interactive control means, wherein the interactive control means puts series of interactive sequences having calling relations together in a plurality of interactive tasks in association with relations and includes an interactive task hierarchical data base for storing the interactive tasks in a hierarchical structure.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a view showing a configuration of an interactive speech interface unit according to a first embodiment of the invention;

Fig. 2 is a view showing a configuration of an interactive task hierarchical data base;

Fig. 3 is a view showing a configuration of an interactive task;

Fig. 4 is a view showing an example of an upper/lower interactive task chain fetched from the data base;

Fig. 5 is a view showing a configuration of an interactive sequence;

Fig. 6 is a view showing an example of an interactive sequence;

Fig. 7 is a view showing an example of storage of an interactive sequence in an interactive sequence storage part;

Fig. 8 is a view showing processing flow by an interactive controller;

Fig. 9 is a view showing a configuration of an interactive speech

3

interface unit according to a second embodiment of the invention;

Fig. 10 is a view showing an example of rewrite of an interactive task chain;

Fig. 11 is a view showing an example of an interactive sequence;

Fig. 12 is a view showing a configuration of an interactive speech interface unit according to a third embodiment of the invention;

Fig. 13 is a view showing a configuration of a user catalog interactive sequence;

Fig. 14 is a view showing an example (1) of a user interactive sequence cataloged dialog;

Fig. 15 is a view showing an interactive sequence cataloged in Fig. 14;

Fig. 16 is a view showing an example of a keyword cataloged dialog;

Fig. 17 is a view showing an example of a bookmark cataloged dialog;

Fig. 18 is a view showing an example (2) of a user interactive sequence cataloged dialog;

Fig. 19 is a view showing an interactive sequence cataloged in Fig. 18;

Fig. 20 is a view showing a dialog using the user interactive sequence in Fig. 15;

Fig. 21 is a view showing a dialog using the user interactive sequence in Fig. 19;

Fig. 22 is an example of an application operation according to a conventional interactive speech;

Fig. 23 is a view showing an example of a conventional interactive sequence.

PREFERED ENBODIMENT OF THE INVENTION

First Embodiment

Fig. 1 is a view showing an interactive speech interface unit according to a first embodiment of the invention. For technical terms according to the invention, a bundle of a series of interactive sequences having a plurality of calling relations is referred to as "interactive tasks". An interactive sequence which is called from other interactive sequences is referred to as a "sub-interactive sequence".

4

A reference numeral 101 is a speech recognition part for recognizing input speech of a user and converting it into a character string, 102 is a database for speech recognition for storing information to be used for speech recognition, 103 is an input statement analysis part for analyzing the recognized character string and converting it into semantic representation, 104 is an information database for input statement analysis for storing information to be used for input statement analysis, 105 is an interactive controller for controlling the flow of interactive status to execute a dialog with the user and access an application via an application interface part 111, described later, and 106 is an interactive task hierarchical database for storing interactive tasks in a hierarchical structure.

Fig. 2 shows an example of the interactive task hierarchical database. Although one upper interactive task is illustrated in the database in Fig. 2, it is permitted that a plurality of interactive sequences are present as the upper interactive task.

Fig. 3 shows a configuration of an interactive task. The interactive task comprises a series of interactive sequences and an interactive sequence prepared by modifying an upper interactive sequence. For example, if a modified version of an interactive procedure of "application operation interactive task" in an initial status is added to "Chinese restaurant retrieval interactive task", it is possible to output "Chinese restaurant retrieval" to the user by speech when starting a dialog.

In the database, the lower interactive task is prepared to include all the sub-interactive sequences which are needed for the upper interactive task. Even if there are a plurality of lower interactive tasks, all the sub-interactive sequences of the upper interactive task need be included in the respective lower interactive tasks.

Fig. 4 shows an example of an upper/lower task chain fetched from the interactive hierarchical database shown in Fig. 2.

A reference numeral 107 in Fig. 1 is an interactive sequence storage part storing interactive sequences included in the interactive task chain fetched from the interactive hierarchical database. With data of the interactive hierarchical database, the interactive task chain as shown in Fig. 4 is first fetched, then it is stored in the interactive sequence storage part while reflecting "modified portion of an interactive sequence of the upper interactive task" shown in Fig 3.

Fig. 5 is a configuration of the interactive sequence. It is assumed that an interactive status name is unique as a whole of the interactive task hierarchy. Further, it is assumed that an interactive status at the transition destination is always present, and the interactive status is always transitive to an end status from any interactive status by applying an appropriate event sequences. At least one interactive procedure corresponds to one interactive status. Although the interactive procedure is written in a programming language and the like, it is written in Japanese language for convenience of the following explanation.

Fig. 6 shows an example of the interactive sequence.

A reference numeral 108 is an output statement generation part for generating an intermediate language to be outputted to a user, 109 is information data base for generating an output statement for use in generating the output statement, 110 is a speech synthesis or generation part for converting the intermediate language into speech, 111 is an application interface part for accessing an application 112 using semantic representation delivered from the interactive controller 105.

The interactive controller 105 fetches an upper/lower chain of the interactive task from the interactive task hierarchical data base 106, converting it into an executable interactive sequence and storing the executable interactive sequence in an interactive sequence storage part 107. It is previously specified as to which interactive sequence chain is fetched when the system is activated.

Fig. 7 is an example of storage of an interactive sequence. A left side in Fig. 7 is an interactive sequence before it is stored in the interactive sequence storage part 107, and a right side in Fig 7 is an interactive sequence after it was stored in the interactive sequence storage part 107. The portions emphasized and described in a boldface correspond to a lower interactive sequence while the portions described in a normal face correspond to an upper interactive sequence. There is "information retrieval interactive sequence" at the modified portion of the upper task of the interactive sequence "restaurant retrieval interactive sequence" before it is stored in the interactive sequence storage part 107. Accordingly, an interactive procedure PROC_101 of the upper "information retrieval interactive sequence" in an interactive status STATUS 101 is replaced by a lower PROC 103.

6

Fig. 8 shows a processing flow by the interactive controller 105. The interactive status is first initialized. Thereafter, an interactive procedure which is applied to the interactive status is executed, and the interactive status is rendered transitive while performing a dialog between a user and an application. Every time one interactive procedure is executed, an input from the user, a response from the application, and an event such as various conditions and the like are checked. An output to the user is executed as an action during the interactive procedure. If a sub-interactive sequence is activated during the interactive procedure, the control is shifted to the sub-interactive sequence so that the sub-interactive sequence is executed in accordance with the processing flow shown in Fig. 8.

When the input speech by the use is fetched, firstly the speech recognition part 101 recognizes user utterance and coverts it into a character string. Then the input statement analysis part 103 analyzes the character string outputted from the speech recognition part 101, converts it into semantic representation used by the interactive controller 105 and delivers it to the interactive controller 105. If an output is needed during the execution of the interactive procedure, the semantic representation is supplied to the output statement generation part 108, then it is converted into an intermediate language. Thereafter, the speech generation part 110 converts the intermediate language into speech and outputs it to the user. If an access to the application is needed during the interactive procedure, it is performed via the application interface part 111. Even in this case, using the semantic representation, the application interface part 111 converts the semantic representation while the application 112 receives a command from the application interface parts 111 and outputs processing results to the application interface part 111.

Second Embodiment

Fig. 9 is a view showing a configuration of an interactive speech interface unit according to a second embodiment of the invention. The second embodiment is different from the first embodiment in respect of the addition of an interactive task chain part 206. The interactive task chain part 206 fetches an interactive upper/lower chain of an interactive task, from an interactive task hierarchical data base 207, during the execution of the dialog, in the same manner as shown in Fig. 4 of the first embodiment of the invention, and rewrites in part.

Only the operation of the interactive task chain part 206 shown in Fig. 9 is different from the operation of the first embodiment.

First of all, in the same manner as the first embodiment, an interactive controller 205 fetches an interactive task chain from the interactive task hierarchical data base 207, converts it into an executable interactive sequence and stores the executable interactive sequence in an interactive sequence storage part 208. It is previously specified as to which interactive sequence chain is fetched when the system is activated. It is different from the first embodiment in that another interactive task chain is fetched from the interactive task hierarchy during the execution of the dialog so that a part of the interactive sequence stored in the interactive sequence storage part 107 is rewritable.

Fig. 10 is an example of rewriting of the interactive task chain. Further, it is possible to rewrite the interactive sequence by describing a processing to restore another interactive task chain in the interactive sequence storage part during the execution of the dialog. The rewriting is described as an action of the interactive procedure. However, it is assumed that the change of an interactive sequence which is under execution and the change of an interactive sequence which causes the change of the portion of an interactive sequence of a calling part is not permitted.

Fig. 11 is an example of description of an interactive sequence for causing rewriting the interactive sequence. An action which is placed first in an initial status is executed when a control is shifted from the application operation interactive sequence to the information retrieval interactive sequence so that the interactive controller supplies a semantic representation that "what do you retrieve?" to the output statement generation part 209. Thereafter, a speech is outputted to the user through the same processing as the first embodiment.

When the user inputs a speech of "French food" and the like, a semantic representation of "French restaurant" is supplied from the input statement analysis part 203 to the interactive controller 205 through the same processing as the first embodiment. A processing to normalize an input of "France" to a representation which is described in the interactive sequence such as "French restaurant" is effected by the input statement analysis part 203. In Fig. 11, a processing is continued after rewriting an interactive sequence which is specified by the action of the rewiring

interactive sequence. When rewriting, if the modified portion of the interactive sequence under execution is included in the lower interactive task, the interactive sequence under execution is rewritten, and hence it is assumed that the rewriting using such an interactive task is not described.

5   Third Embodiment

Fig. 12 is a view showing a configuration of an interactive speech interface unit according to a third embodiment of the invention. The third embodiment of the invention is different from the first embodiment of the invention in respect of constituents 307, 308, 309 and 310. According to 10 the third embodiment, a user can catalog an interactive sequence. A character string which the speech recognition part recognized as well as semantic representation are delivered from the input statement analysis part 303 to the interactive controller 305. Further, when the keyword is cataloged, the interactive controller 305 can use the speech recognition 15 character string as an event as it is. The bookmark is a name attached by a user for an interactive status and it is used for specifying the interactive status of a destination of transition when the interactive sequence is cataloged by the user.

A reference numeral 307 is a keyword/bookmark catalog interactive 20 sequence storage part for storing an interactive sequence for use in cataloging a keyword and a bookmark, 308 is a keyword/bookmark storage part for storing the keyword and the bookmark specified by the user, 309 is a user interactive sequence catalog interactive sequence storage part for storing an interactive sequence for use in cataloging the user interactive 25 sequence, and 310 is a user interactive sequence storage part for storing the user interactive sequence .

Fig. 13 is a configuration of the user catalog interactive sequence used in the third embodiment. The third embodiment is different from the first embodiment in that a keyword can be used as an event.

30   The operation of the third embodiment which is different from that of the first embodiment is now described. First of all, an operation for cataloging a user interactive sequence is described. Fig. 14 is an example (1) of a catalog of a user interactive sequence. The user moves to an interactive status where an interactive procedure is to be added. When 35 moved to the interactive status, if there occurs a transition to another status, an interactive sequence stored in the user interactive sequence catalog

9

interactive sequence storage part is activated by a previously specified input. The user specifies an event, an action and an interactive status to be transitive next.

For the event, events used in that status are read out by the system, and one of the events is to be selected. In addition to that, "no-event (unconditional)" and a keyword (setting manner is described later) can be used.

For the action, actions capable of being used in that status are read out by the system, and one of the actions is to be selected. In addition to that, "no-action (nothing is done)" can be selected.

A next transition status is specified by use of a bookmark cataloged by a user (descriptive manner is described later). The user interactive sequence is stored in the user interactive sequence storage part 310. Fig. 15 is a user interactive sequence cataloged in Fig. 14. The keyword and the bookmark are cataloged in the following manner.

(1) Since the keyword and the bookmark are cataloged in correspondence with the interactive status, they are first moved to the interactive status so as to correspond thereto.

(2) The interactive sequence stored in the keyword/bookmark catalog interactive sequence storage part 307 is activated. A user catalogs the keyword and the bookmark in the specified manner (see Figs. 16 and 17). In an "end status", if transit to that status, the dialog system per se is ended, and hence the bookmark is not attached so that it is necessary that the system can prepare and specify the bookmark. Except for that, in an interactive status where the bookmark can not be attached, the system prepares the bookmark. The reference of inspection of release of the keyword and the bookmark is processed in the similar procedure. In a status such as an end status where a user can not actually attach the bookmark in that status, it can be used as a reserved word. The release of the bookmark is prohibited in the similar interactive status by the catalog of the user interactive sequence described later.

For the event of the user interactive sequence, a cataloged keyword can be used. Fig. 18 shows an example (2) of the catalog of the user interactive sequence using the keyword. Fig. 19 is an interactive sequence cataloged in Fig. 18. It is assumed that in an interactive status where the user can not stay by the occurrence of any action, an interactive sequence

10

capable of cataloging a status from other interactive statuses while specifying a status name is described in a user interactive sequence catalog dialog. The reference of inspection of deletion of the user interactive sequence is also processed in the similar procedure.

5   The operation of the interactive speech using the user interactive sequence is described next. A character string recognized by a speech recognition part 301 is added to semantic representation which is delivered from the input statement analysis part 303 to the interactive controller 305. The interactive controller 305 retrieves the keyword/bookmark storage part 10 to check as to whether the character string attached to the semantic representation is the keyword which corresponds to the present interactive status. If cataloged, the keyword is managed as an event. If not cataloged, the semantic representation is managed as an event.

For a method of application of the interactive sequence, an 15 interactive sequence stored in the user interactive sequence storage part 310 is first applied. If the interactive procedure to be applied is not found, an interactive sequence stored in the interactive sequence storage part 311 is applied.

Other operations are the same as those of the first embodiment. 20 Fig. 20 is a dialog using the user interactive sequence shown in Fig. 15 and Fig. 21 is a dialog using the user interactive sequence shown in Fig. 19.